Foundations of Natural Language Processing
Lecture 19c
Constructing Representations of Discourse Coherence

Alex Lascarides

So Far

- Understanding discourse involves identifying the coherence relations
  - Narration, Explanation, Background, Contrast, Parallel, QA, Correction...
- Discourse coherence influences many pragmatic phenomena.
- LFs that feature coherence relations can be formally interpreted

Now: How to automatically infer coherence relations

Recap: an Example

π_1: John can open Bill's safe.
π_2: He knows the combination.

$$\begin{array}{ll} \pi_0: & \textit{Explanation}(\pi_1, \pi_2) \\ \pi_1: & \iota x(\textit{safe}(x) \wedge \textit{possess}(x, \textit{bill}) \wedge \textit{can}(\textit{open}(e_1, \textit{john}, x)) \\ \pi_2: & \iota y(\textit{combination}(y) \wedge \textit{of}(y, x) \wedge \textit{knows}(\textit{john}, y)) \end{array}$$

- Bits in red are specific values that go beyond content that's revealed by linguistic form.
- They are inferred via commonsense reasoning that's used to construct a maximally coherent interpretation.

Symbolic approaches to constructing LF

- Draw on rich information sources:
  - linguistic content, world knowledge, mental states...
- Deploy reasoning that supports inference with partial information. Unlike classical logic, this requires consistency tests.
- Typically, construct LF and evaluate it in the same logic, making constructing LF undecidable.

Further Problem

- Like any knowledge rich approach involving hand-crafted rules, this is only feasible for very small domains.

- Ideally, we would like to learn a discourse parser automatically from corpus data.

- But there's a lack of corpora annotated with discourse structure.
  - RSTbank, Graphbank, Annodis, STAC are relatively small.
  - Discourse Penn Treebank is relatively large but not annotated with complete discourse structure.
  - Groningen Parellel Meaning Bank: full discourse structure (SDRSs) and getting bigger all the time.

Supervised Learning for SDRT

Training on 100 dialogues — Baldridge and Lascarides (2005)
Parser based on Collins' parsing model:

- 72% f-score on segmentation (baseline: 53.3%)
- 48% f-score on segmentation and coherence relations (baseline: 7.4%)
- Doesn't attempt to estimate LFs of clauses.

Training on Groningen Meaning Bank — Liu and Lapata (2018)
Neural semantic parser, RNN computes structure first, fills in arguments later:

- 77% f-score on segmentation, coherence relations *and* LFs of clauses
- State of the Art!

Avoiding Annotation

Sporleder and Lascarides (2008)

- Coherence relations can be overtly signalled:
  - *because* signals EXPLANATION; *but* signals CONTRAST

- So produce a training set *automatically*:
  - Max fell because John pushed him
    $\Rightarrow$
    EXPLANATION(*Max fell*, *John pushed him*).

Results of Best Model

- Test examples originally had a cue phrase: 60.9%.
- Test examples originally had no cue phrase: 25.8%
- Train on 1K manually labelled examples: 40.3%.
- Combined training set of manual and automatically labelled examples doesn't improve accuracy.

So you're better off manually labelling a small set of examples!

Results of Best Model

- Test examples originally had a cue phrase: 60.9%.
- Test examples originally had no cue phrase: 25.8%
- Train on 1K manually labelled examples: 40.3%.
- Combined training set of manual and automatically labelled examples doesn't improve accuracy.

So you're better off manually labelling a small set of examples!

Why?

Contrast to *Elaboration*
Although the electronics industry has changed greatly, possibly the greatest change is that very little component level manufacture is done in this country.

Summary

- Interpretation governed by discourse coherence:
  - Constrains what can be said next
  - Augments meaning revealed by linguistic form.

- Computing logical form should be decidable; modularity is key to this.

- Data-driven approaches are a major challenge.

- Linking rich models of discourse semantics to models of human behaviour and decision making is also a major challenge, but essential for tackling dialogues where the agents' goals conflict.